电视音频技术发展现状和展望

摘　要：随着电视视频技术的不断发展，电视音频也从单声道、双声道发展到多声道及 3D 环绕声的形式。本文从音频技术高清化演进路线入手，介绍了当今主流的多声道及 3D 环绕声格式，并对元数据的控制功能进行阐述。通过对现有技术的梳理，预测了电视音频技术会向着全维度再现发展，并且随着元数据功能的完善，未来的电视音频会具有更好的交互性及灵活性。

关键词：高清时代；3D 环绕声；元数据；全维度
中图分类号：G220.7　　**文献标识码：**A
文章编号：1671-0134（2017）12-061-04　　**DOI：**10.19483/j.cnki.11-4653/n.2017.02.011

■文 / 张　磊　姜世杰

自从 20 世纪人类发明黑白电视以来，电视已经成为信息采集、交换和传播的主流媒体。电视技术已经经历三次质的飞跃，20 世纪 50 年代电视从黑白过渡到彩色，给观众带来了前所未有的彩色视觉体验；20 世纪 90 年代随着数字技术和网络技术的发展，电视从模拟的单一传像功能向数字化的智能型、交互式、多用途方式转变；而第三代电视是数字高清晰度电视和 3D 立体电视，带给观众临场感极强的高清晰度视频图像和极具震撼力的音响效果。在模拟电视时代，电视系统最多只能提供两个声道的链路。当电视的传输链路从模拟进入数字时代以后，电视伴音也从单声道过渡到双声道和多声道。尤其当数字电视的高清时代到来以后，多声道的环绕声系统已经成为一种必然。

1. 音频技术高清化的演进路线

人类对于电视伴音重放质量的要求总是在随着科技水平的提高而不断增长。而这其中最显著的变化，就体现在重放系统声道数量的增加上。人类的听觉是一种对多种信息的综合性反应。除了最基本的声音要素，如响度、音调和音色以外，还包括反映声音空间特性的因素，比如声源的方向、声源的远近、声场的大小、声场的色彩等。而电视伴音系统声道数的不断增长，也就是为了能够将这些信息尽可能多地予以体现。

在模拟电视时代，电视伴音采用单声道系统，只使用一只或几只扬声器来重放由一只传声器记录的声音信号。这种重放系统只能表现声音的响度、音调和音色，并在一定程度上体现声源的远近，而对于声源的空间定位等重要特性并不能表现出来。单声道的电视伴音为模拟信号，采用调频传输，这种信号在在传输、存储和变换过程中常会产生下列问题：

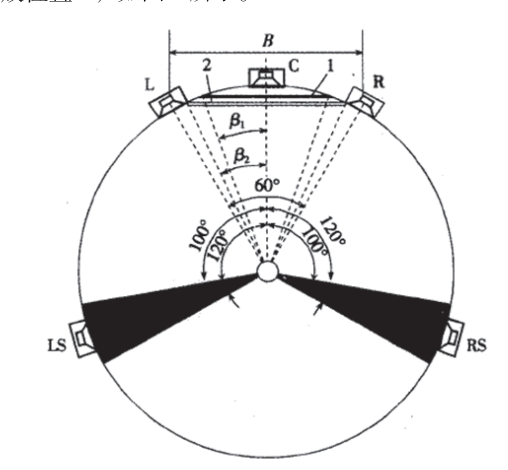
①音频信号存储载体的信号动态范围只有 40 ～ 50 dB，远低于节目源的最大信号动态范围（120 dB）；

②在信号编辑和变换中（节目编辑、转录和延时效果处理等）随着变换次数的增加，音质会迅速恶化。

因此，为了改善模拟单声道电视伴音质量，音频的数字化发展成为必然趋势。进入数字电视时代，针对标清数字电视，电视伴音为单声道或双声道立体声数字信号；针对高清数字电视，电视伴音为多声道数字环绕声信号。音频信号的数字化，很容易实现大于 90dB 的动态范围。此外数字音频信号可以进行非线性编辑，而不会增加音频信号的失真。

对于标清数字电视的立体声双声道系统，能够利用“双耳效应”来产生“听觉幻像”，实现对前方声源横向、纵向的定位，并产生比较明显的表现声音空间特性。虽然该重放系统相对模拟单声道时代在音质和听音效果上有了较大提升，但是与实际的声场还有很大的差距，其主要问题是听者的侧方和后方区域声场没有如实地反映出来。

进入高清时代以后，随着电视画面的清晰度越来越高，为了得到更佳的视听享受，多声道环绕声系统的电视伴音成为必然。多声道环绕声系统追求的是对声音空间特性的全方位表现，给人们三维的立体空间印象。目前数字电视音频标准中都采用 5.1 环绕声系统，也就是 3/2/.1 的配置方案。这种方法是按照 ITU-R BS.775 的建议来确定用于重放的扬声器摆放位置 [1]，如图 1 所示。



扬声器	水平角度 (°)	高度（m）	垂直角度 (°)
C	0	1.2	0
L R	30	1.2	0
LS RS	100 ～ 120	≥ 1.2	向下倾斜 0 ～ 15

图 1　ITU-R 建议的 5.1 声道扬声器设置

人们为了获得更稳定的声像定位和覆盖范围更大的听音

区域，多声道环绕声系统已经由 5.1 声道扩展为 7.1、10.2，甚至 22.2 声道系统，极大地丰富了声音的再现能力。目前 10.2 和 22.2 声道的环绕声系统还处于体验阶段，而 7.1 声道的环绕声系统已经成为绝大多数高清影片的伴音格式，常用的扬声器摆位如图 2 所示。7.1 声道环绕声系统增加了一组环绕声扬声器(Lb 和 Rb)，而将 5.1 声道系统的环绕扬声器(Ls 和 Rs) 前置，将环绕声场分解成前侧方 (30° ~90°)、后侧方 (90° ~150°) 和后方 (150° ~180°) 三个部分，进一步增强了环绕声场声像定位的连续性和包围感。

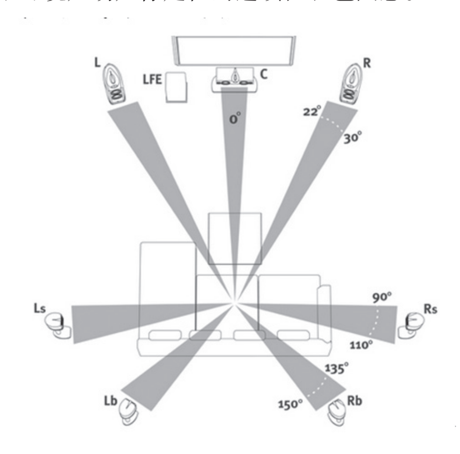


图 2　7.1 声道环绕声系统扬声器摆位

与立体声双声道系统对比，多声道环绕声系统存在以下优势：

①多声道环绕声由于添加了侧后方扬声器，可以在一定程度上实现对后方声源的重放。

②多声道系统对声源方向的安排更加灵活。在奥运会转播过程中，评论员的声音由中置声道重放，运动声响主要由左、右声道重放，观众的欢呼声及运动场环境声则由左、右、左环绕和右环绕声道重放，可以较好地将不同声音元素分离开 [2]。

③相比于双声道立体声，多声道环绕声拓展了聆听区域。双声道立体声系统，要求听者必须位于距两扬声器等距离的某一点，才能获得比较满意的声像感。而多声道环绕声系统的听音范围比较宽，处于环绕声系统最佳听音位置附近的听者对声像的感受不会产生很大差异，有利于多人一起欣赏。

④运用多声道环绕声系统能够还原声源所在声场的声学特点，即空间感、包围感、温暖感等。人们在观赏多声道环绕声的音乐会转播时，就好像在真实的音乐厅中聆听现场演奏一般。

2. 高清时代的多声道环绕声格式

多声道环绕声格式在 20 世纪 90 年代初开始逐渐普及，到现在为止已经有十几种多声道环绕声格式，下面按照声道数的递增，选取典型的环绕声格式进行梳理和阐述。

2.1　5.1 声道环绕声格式

目前 5.1 声道的环绕声格式是最为普及的系统，常见的环绕声格式包括 Dolby Digital、DTS 等。Dolby Digital 格式由 Dolby 公司开发 [3]，主要应用于专业电影、广播电视和家庭影院。该格式采用 AC-3 编解码技术，压缩率可达 10:1，支持的最高数据传输率为 640kbit/s。目前在广播电视领域，多数电视台采用 Dolby Digital 和 Dolby E 技术进行数字高清电视节目的录制和传输，如图 3 所示。在电视制作端采用 Dolby E 技术 [4]，该技术可以通过一个 AES 数字音频对传输多达 8 声道的数字音频信号。目前电视广播基础设施 (录像机) 绝大部分只有 2 声道或 4 声道的音频处理能力，采用 Dolby E 技术可以将多声道信号存储在现有的设备上，不用更新设备即可传输多声道音频信号。此外经过 Dolby E 技术编码的音频帧时长为 40ms，与视频帧相同便于进行声画的同步编辑。在电视播出前用 Dolby Digital 编码将多种格式(5.1 声道、单声道、双声道立体声) 数字音频信号进行压缩编码为一对 Dolby Digital 码流，输入 MPEQ 编码 / 复用。在用户接收端，利用机顶盒对 Dolby Digital 码流进行解码，还原编码前的音频信号，输入到用户的家庭影院。

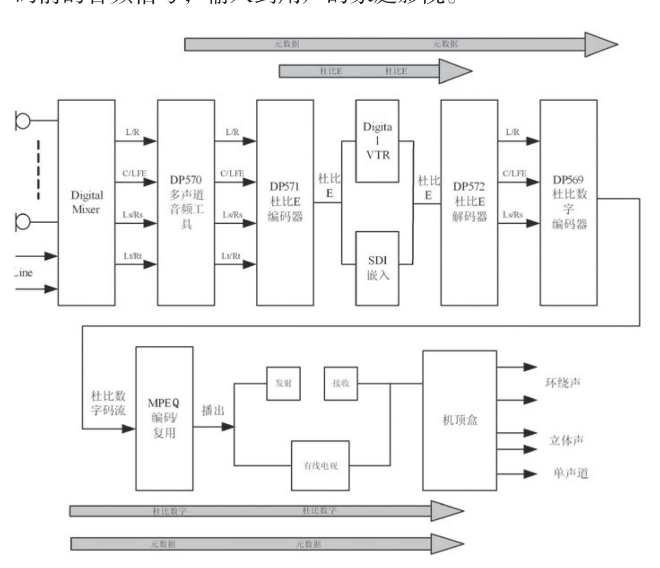


图 3　Dolby 公司推出的数字电视音频解决方案

DTS 格式由 DTS 公司开发，主要应用在专业电影、家庭影院和纯音乐领域中。该格式采用相干声学编解码技术，压缩率在 2.9:1 到 4.3:1 之间，取样率在 8 ~ 192kHz 之间，量化精度在 16 ~ 24bit 之间。在 1993 年 6 月 11 日放映的《侏罗纪公园》首次采用了该系统，该系统使用了声画分离的方式，DTS 码流记录在 CD-ROM 上。因此只要将录制数字音频信号的 CD 单独拿出来播放，就能得到多声道环绕声音效。如果将多声道音乐按 DTS 格式录制在一张 CD 上，用普通的 CD 机播放，并在其数字输出口处接上一台 DTS 解码器，就可以得到环绕声音乐。这种光盘被称为 DTS-CD。

2.2　7.1 声道环绕声格式

在电影领域，随着 BD 光盘的普及，高清电影已经成为主流形式。各大公司都为在高清领域争得一席之地而努力研发自己的新型环绕声系统。杜比公司就在 2004 年的东京 AES 大会上，首次展示了全新的 Dolby Digital Plus 系统 [5]，随后还推出了 Dolby True HD 系统 [7]，这两个系统都是为高清光盘格式的发展而设计的多声道音频格式。而 DTS 公司也不甘示弱，推出了 DTS HD 系统 [6]。目前这三款音频格式都已经纳入 BD 光盘的音频标准格式中。纵观这几种新型的系统，它们都以 7.1 声道为起点，可向上扩展。此外声音的

再现也不仅仅局限在水平维度上，通过扬声器的配置可再现垂直维度的定位感，极大增强了声音的表现力。

2.3 加入高度声道的 3D 环绕声格式

当前人们对声音的再现能力开始追求“全息立体声”的效果。SMPTE 成立的 DC28 数字电影技术委员会给未来数字电影制定了可容纳声道数量与扬声器的配置方式（《SMPTE 428.3M 协议》），如图 4 所示。该配置方式描述了 20 个声道的设定，除了水平维度设置的 16 个声道，还有 4 个声道当作垂直阵列，用于增强高度层次感。

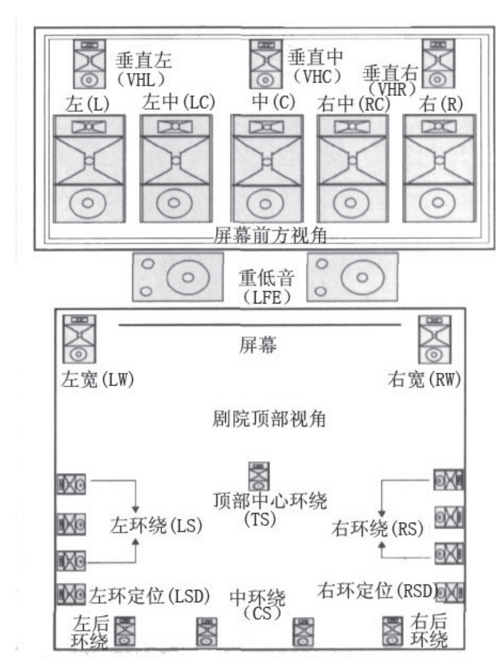


图 4　SMPTE 428.3M 的扬声器配置及分布

日本 NHK 公司为了配合超高清电视 (UHDTV) 而推出了 22.2 声道的环绕声系统 [7]，如图 5 所示。整个系统有 10 只扬声器位于听者头部所在的水平面，9 只扬声器高于听者头部水平面，其他 3 只扬声器和 2 只重低音扬声器设置在低于听者头部水平面。虽然该系统仍在推广体验阶段，但是由此可以看出数字电视的音频系统将朝着多通道的全维度方向发展。

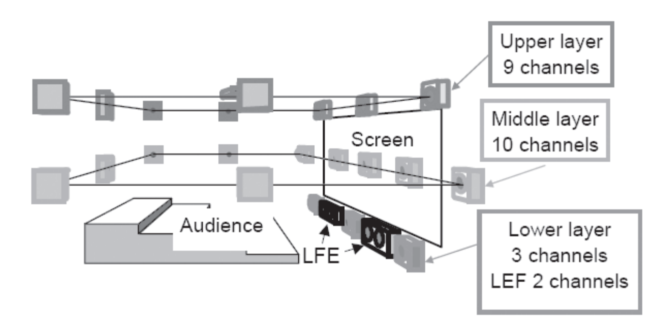


图 5　NHK 针对超高清电视推出的 22.2 声道环绕声系统

Dolby、DTS 和 Baccro 等公司近几年也分别推出了 Dobly Atmos、DTS:X 和 Auro-3D 等 3D 环绕声重放系统。Dolby Atmos 及 DTS:X 针对家庭影院推出了 7.1.4，7.1.2，5.1.4，5.1.2 等四种重放格式，图 6 显示了 DTS:X 的 7.1.4 重放系统。Auro-3D 的家庭影院系统包括 9.1 和 10.1 两种格式，图 7 显示了 Auro 3D 的 10.1 格式重放系统。由于高度声道的增加，可以增强空间声源的定位及有效的扩展三维空间感，为观众带来更加真实的体验。



图 6　DTS:X 的 7.1.4 重放示意图

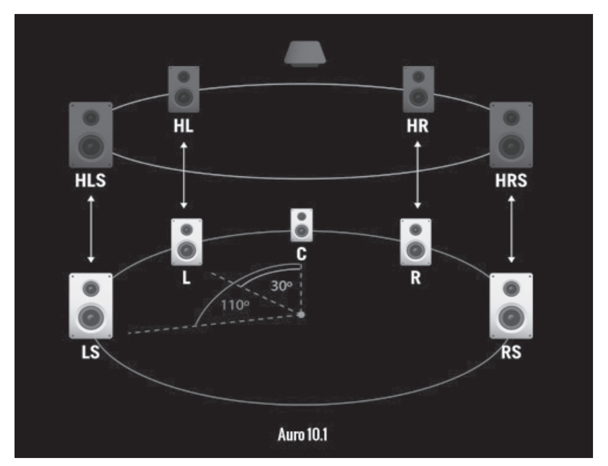


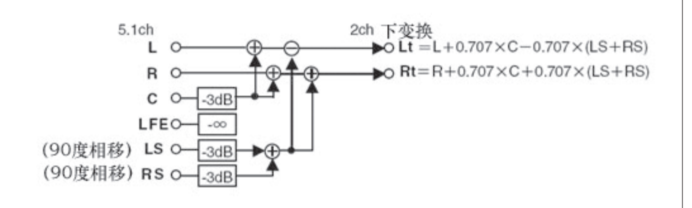
图 7　Auro 3D 的 10.1 重放系统示意图

3. 多声道格式中元数据的控制功能

为保证数字电视音频信号的正确传输和接收，需要使用元数据作为贯穿节目从制作到播出和接收完整链路的控制手段。所谓元数据，其本质的意义是关于音频数据的数据，即一些音频描述和控制参数，例如下混合参数和对白归一化参数等。不同的多声道编解码技术都采用元数据来进行音频信号的控制。

3.1 下混合元数据

下混合元数据是为使多声道节目能够被立体声和单声道用户收看收听而设定的。由于低音增强 LFE 声道记录的音频信号主要用于渲染烘托气氛，所以在多声道节目下变换成双声道或单声道时，只用其中的 L、R、C、Ls、Rs 五个主声道。图 8 和图 9 分别显示了 Dolby Digital 和 MPEG 2-AAC[8] 提供的多声道节目下变换到双声道的算法。



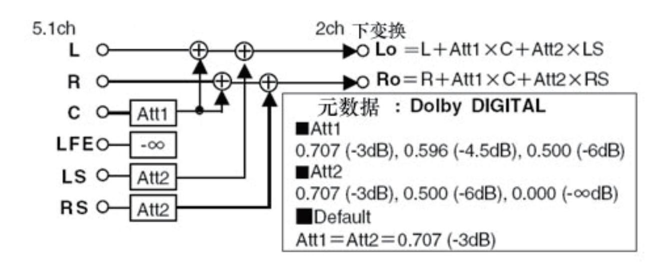
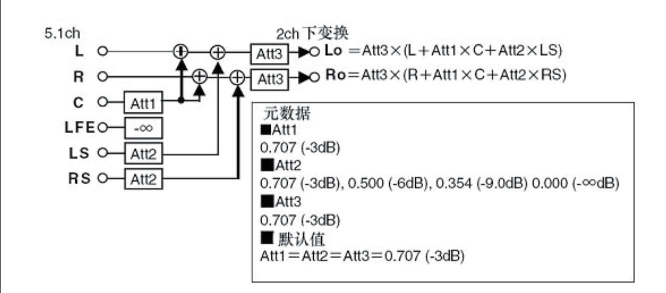


图 8 Dolby Digital 提供的两种下变换算法



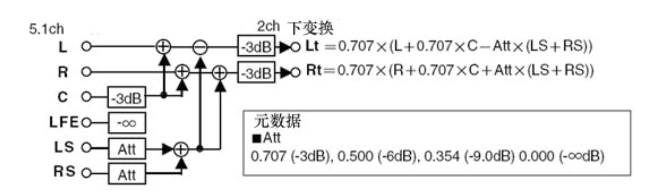


图 9 MPEG 2-AAC 提供的两种下变换算法

3.2 对白归一化

当重放不同来源的音频时，进行节目切换时常常出现响度不一致的情况。播放的音频可能是同一频道不同的节目类型（新闻，广告），也可能是不同的播放频道。为了解决这个问题，在进行编码时往往将参考电平直接编码到音频数据流中。通常使用正常口头对白的主观声压级作为参考电平。

在解码端，参考电平值被用在重放系统中设定重放音量，对系统的音量控制一般根据听众期望的响度或根据声压级（SPL）来设定。应用参考电平值，主要包含以下的功能：达到听众期望的对白声压级，以及指出在音频信号中对白声压级的参考电平值。这样不论被解码的什么类型的节目，听众都能够设定对白音量，让对白的主观响度保持一致。

如图 10 所示几种不同的音频信号，其平均对白电平如图数字标记所示。在切换不同节目时，由于电平的不一致，会导致不同的主观声压级。如果将电影的对白电平降低 3dB 到 -30dB 作为主观声级，即参考电平值，通过应用参考电平值，再切换成不同类型节目时，都会自动的将节目中的对白电平归一化为 -30dB，以达到一致的相同主观感知，如图 11 所示。

4. 结论

模拟电视时代人们对声音在电视系统中所起的作用有所忽视，一直将其称为电视伴音。随着电视从模拟标清发展到数字高清晰度电视、超高清晰度电视以及立体电视，人们在追求视觉临场感的过程中，音频将对视频起到极大的辅助作用。声音从单声道、双声道立体声发展到多声道环绕声格式已经成为一种必然趋势，相信随着数字电视音

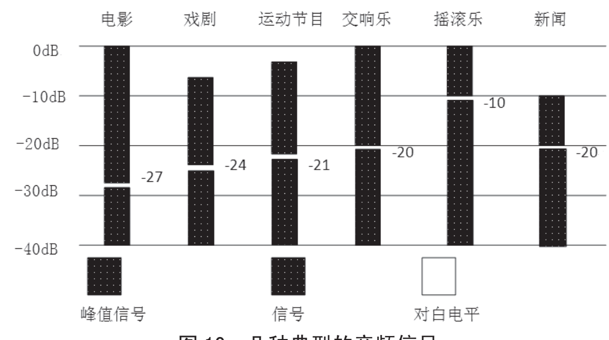


图 10 几种典型的音频信号

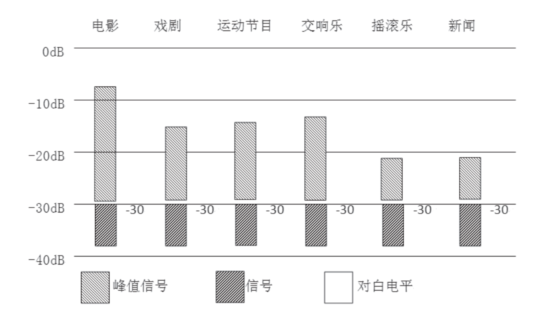


图 11 归一化后的音频信号

频技术的进一步发展，声音将实现全维度的再现。而随着元数据功能越来越完善，未来的电视音频技术会增加更多的互动性及灵活性。媒

参考文献

[1] ITU-R Recommendation BS.775-1. Multichannel Stereophonic Sound System with and without Accompanying Picture. International Telecommunication Union, Geneva, Switzerland, 1992-1994.

[2] 王树森 . 奥运转播国际公用信号制作的音频技术应用 [J]. 电声技术，2008，32（5）：4-10.

[3] ATSC A/52B：2010. Digital Audio Compression Standard（AC-3, E-AC-3）. Advanced Television Systems Committee, Washington, D.C., 22 November 2010.

[4] 孔晓蕾 . 杜比 E 技术简介 . 电声技术，2003（5）：61-63.

[5] Roger Dressler. Dolby Audio Coding for Future Entertainment Formats. Dolby Laboratories, Inc., San Francisco, CA, White Paper, 2006.

[6] DTS-HD Audio Consumer White Paper for Blu-ray Disc and HD DVD Applications. DTS, Inc., White Paper, 2006.

[7] Kimio Hamasaki, Toshiyuki Nishiguchi, Reiko Okumura, Yasushige Nakayama, Akio Ando1. A 22.2 Multichannel Sound System for Ultrahigh-Definition TV（UHDTV）. SMPTE Motion Imaging Journal, 2008, 117: 40 - 49.

[8] ISO/IEC 13818-7. Information technology - Generic coding of moving pictures and associated audio information-Part 7: Advanced Audio Coding（AAC）.

（作者单位：北京电视台）